

AD-A162 146

0

Cost Uncertainty Assessment Methodology:
New Initiatives

by

K. T. Wallenius
Clemson University

Presented at the
19th Annual Department of Defense
Cost Analysis Symposium
Xerox Training Center
Leesburg, Virginia
September 17 - 20 1985

FILE COPY

DTIC
ELECTED
NOV 26 1985
S E D

This document has been approved
for public release and sale. Its
distribution is unlimited.

11 19-85 199

COST UNCERTAINTY ASSESSMENT METHODOLOGY:
NEW INITIATIVES

K. T. Wallenius
Clemson University

OVERVIEW

Accession For	
NTIS GRA&I	
DTIC TAB	
Unannounced	
Justification	
<i>Person 50</i>	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Given the current emphasis on improving all aspects of cost analysis and the primacy of cost uncertainty in the cost estimation process, it is time to assess the quality of the state-of-the-art and look toward improving techniques where needed. An argument supporting the primacy of cost uncertainty analysis and a critical assessment of existing techniques was given in [1]. In this paper, we examine some efforts to apply new results and insights, both mathematical and psychological, to the problems of cost uncertainty analysis. The paper will treat some new ideas pertaining to probability elicitation/encoding using interactive software and indirect assessment of the uncertainty associated with a CER.

1. PROBABILITY ELICITATION/ENCODING

Even in the presence of perfect information and the best application of sound cost analysis techniques, predicting the cost⁶ of, say, an apogee kick motor (AKM) of a proposed unmanned spacecraft entails unknown quantities. One may attempt a bottoms-up engineering approach listing design costs, material and labor requirements, etc., or use a cost estimating relation (CER) developed from data on analogous systems and their costs. Whichever approach is taken, perhaps both, the actual cost is not predictable with certainty. Since the uncertainty associated with AKM cost

⁶ The term "cost" is used in a generic way in this paper. Its meaning will depend on the purpose of the analysis (e.g. first unit cost, nonrecurring cost, life cycle cost, etc.)

will contribute to the overall system cost uncertainty, its contribution must be expressed in some analytically useful form, generally a probability distribution. Using the CER approach, suppose that the costs of AKMs are related to total impulse by the simple linear model

$$SC_i = a + b(\text{impulse})_i + e_i$$

where x is total impulse (measured in lb-sec), SC is standardized cost, a and b are unknown constants, and e represents variability not accounted for by x , often referred to an error or random noise. Thinking of the present system and the historical systems as being a sample from, say, a hypothetical population of all possible AKMs, the parameters a , b , and the variance of the random noise can be estimated in the usual way. For any value x^* of impulse, the fitted model

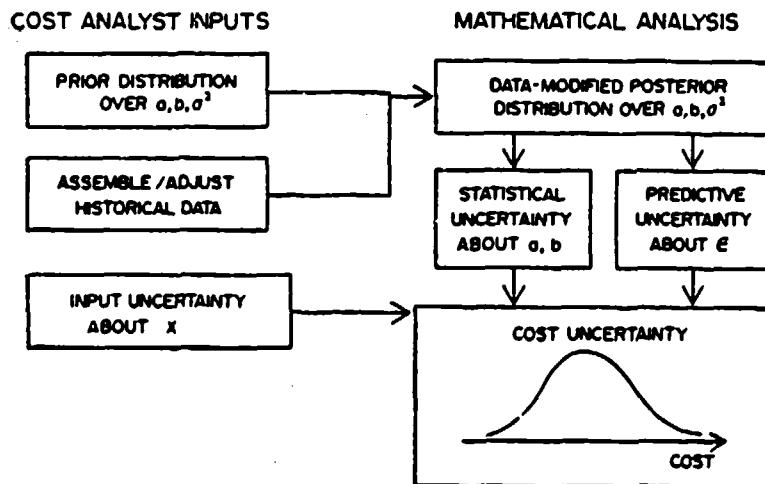
$$\hat{SC} = \hat{a} + \hat{b}(x^*)$$

can be used to get point estimates and prediction intervals for the cost. But what about an uncertainty distribution over the cost? Even if normal theory applies, the predictive distribution of SC is only partially known. It depends on the unknown values of a and b and on the unknown population variance. Even if these values were assumed known, as they are in [7] and [13], there is uncertainty about the value x^* of the input variable x (called strategic uncertainty in [1]). The impulse of the finished product will probably differ from its designed nominal value since the required impulse will depend on the weight of the payload about which there is uncertainty. These various types of uncertainty are discussed in [11]. In order to generate an uncertainty distribution over SC , (called a "predictive distribution" since its purpose is to predict SC), one must quantify uncertainty about a , b , x , and the random noise variance. Direct probability encoding on some or all of these quantities is a difficult cognitive task. New techniques supported by interactive computer codes have been developed [8] which make these tasks easier. They will be discussed in the next section. Shown in figure 1 is a schematic

diagram of how the various types of uncertainty come together to generate the predictive distribution.

FIGURE 1: CER Uncertainties

$$\text{MODEL: COST} = a + b \cdot (x) + \epsilon, \epsilon \sim N(0, \sigma^2)$$



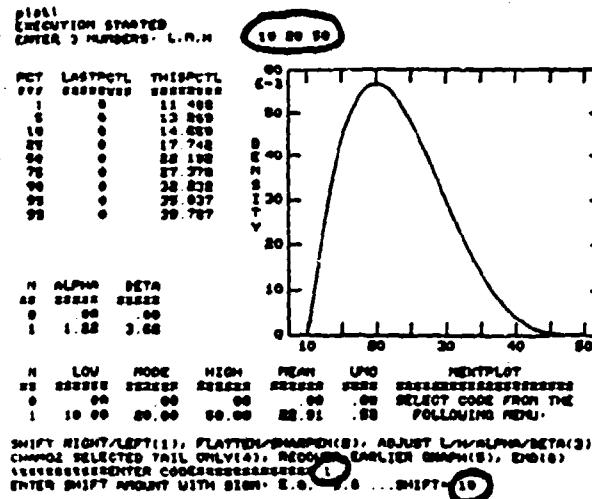
To be specific, what is involved in (direct) encoding of the uncertainty about x , the impulse? The analyst may look at data pertaining to the growth of impulse (i.e. the difference between nominal and actual) over similar AKMs. He may then assess a few key quantities from which a first approximation to the uncertainty distribution is generated. If he is using the standard PERT-beta approach (discussed in [11]) he will assess a lowest possible value (l), a most likely value (m), and a highest possible value (h). A fourth assessment relating to spread (s) must also be given. Various ways for doing this are summarized in [9]. This process can be thought of as specifying a few conditions (points or properties) on the cumulative distribution function (CDF). For example, l and h are the two extreme points, while m specifies the location of the inflection point. Alternately, a few central percentiles might be specified and a CDF "faired"

in or fitted mathematically. Behavioral studies seem to support this latter method for producing better calibrated assessments. See [6] and the cited references for more discussion on the important cognitive issues involved. Whatever method is used, the result is a CDF characterizing uncertainty about the unknown quantity.

Generating an entire CDF based on a few inputs involves considerable extrapolation. The generated CDF is supposed to embody all relevant information including experience-based subjective judgment of the cost analyst. He should carefully examine the implications of the fitted CDF to insure its fidelity and make adjustments where needed. A four-step graphical method was proposed in [9] for accomplishing this type of adjustment. It involves choosing a beta shape from a collection of 9 to 15 graphs. Whenever two consecutive shapes coincide, the analyst is termed consistent and the process stops. The suggested four-step method somewhat resembles an earlier approach [10] which was implemented with a prototype interactive computer program. Because of its popularity, the PERT-beta methodology was chosen for this implementation. Software, using the more versatile and applicable Johnson family of distributions discussed below, is being developed. Crude pictures of successive output screens of the beta code are shown below. User-supplied inputs are circled.

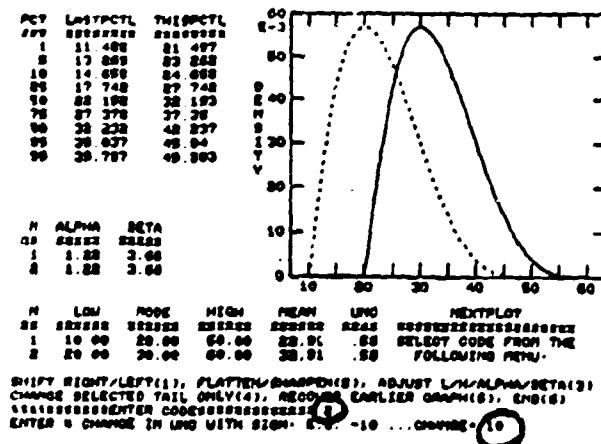
In the first screen, the user is prompted for the standard PERT input: l , m , h . A beta density with the PERT variance is generated. Some selected percentiles are displayed along with the beta shape parameters and other descriptive information (UNC is the "UNCertainty Coefficient" defined in [4]. It is the standard deviation divided by that of a uniform distribution over the same range. Thus, since attention is limited to unimodal betas, $UNC = 1$ indicates a most diffuse opinion. (Smaller values indicate less uncertainty.) A menu of available modifications is printed and the user is prompted to select an option.

FIGURE 2-1: Initial selection



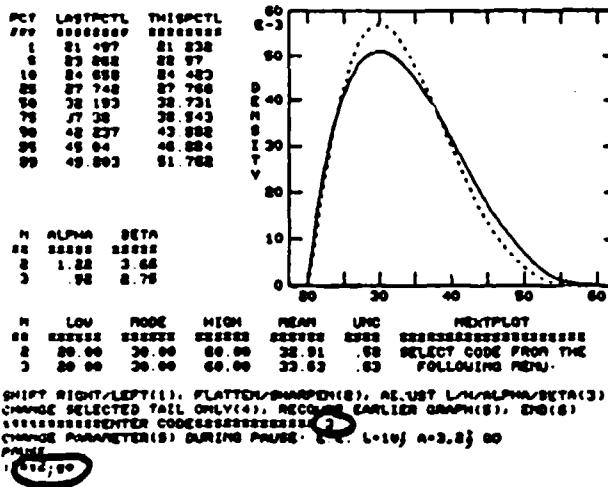
The output of the initial specification is shown above along with the menu. Option 1 is selected and the shift value entered producing

FIGURE 2-2: Result of shifting the density



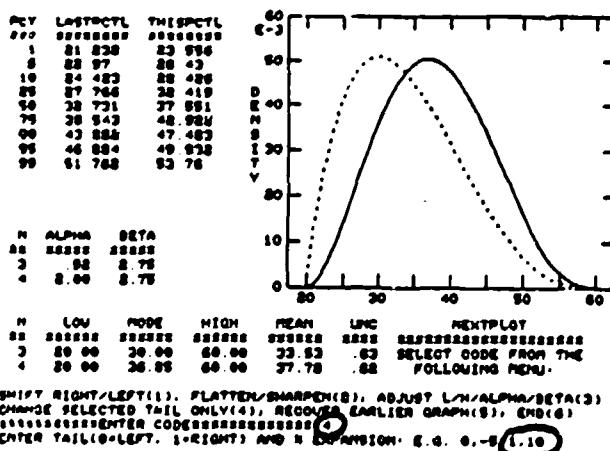
The modified (solid) and previous (dash) densities are shown in figure 2-2. Percentiles and other descriptive information is listed for both densities and the menu repeated. Option 2 (alter the variance but not the range) is chosen and a "flattening" of 10% is specified.

FIGURE 2-3: Result of flattening the density



The flattening effect is evident while l, m, and h are unaltered. Option 3 allows one to set any or all of the parameters. Option 3 is selected and alpha is changed to 2 resulting in

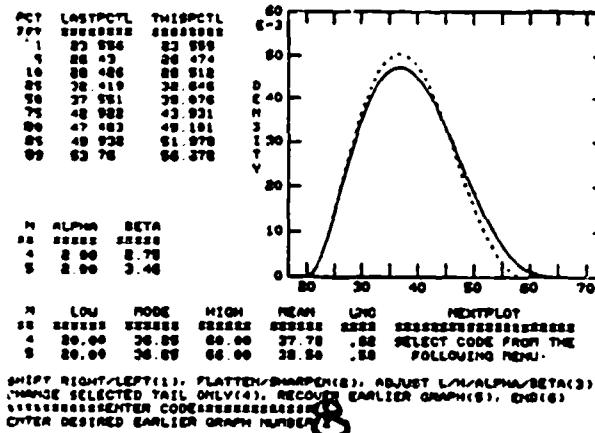
FIGURE 2-4: Result of adjusting parameter(s)



Alpha has been changed from .92 to 2 thus shifting the mode to the right.

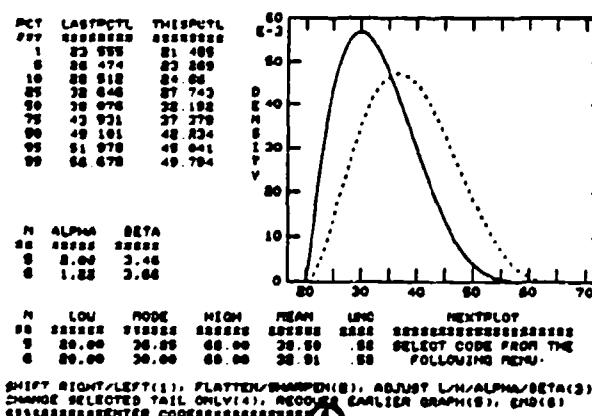
Option 4 allows the user to expand or contract one tail while keeping the mode and other tail fixed. Here, the right tail is expanded by 10%.

FIGURE 2-5: Result of adjusting one tail



The user can recover a previous density by exercising option 5. This saves having to start all over if a mistake is made along the way. Here, the user recovers graph n = 2. It is assigned sequence number 6 in

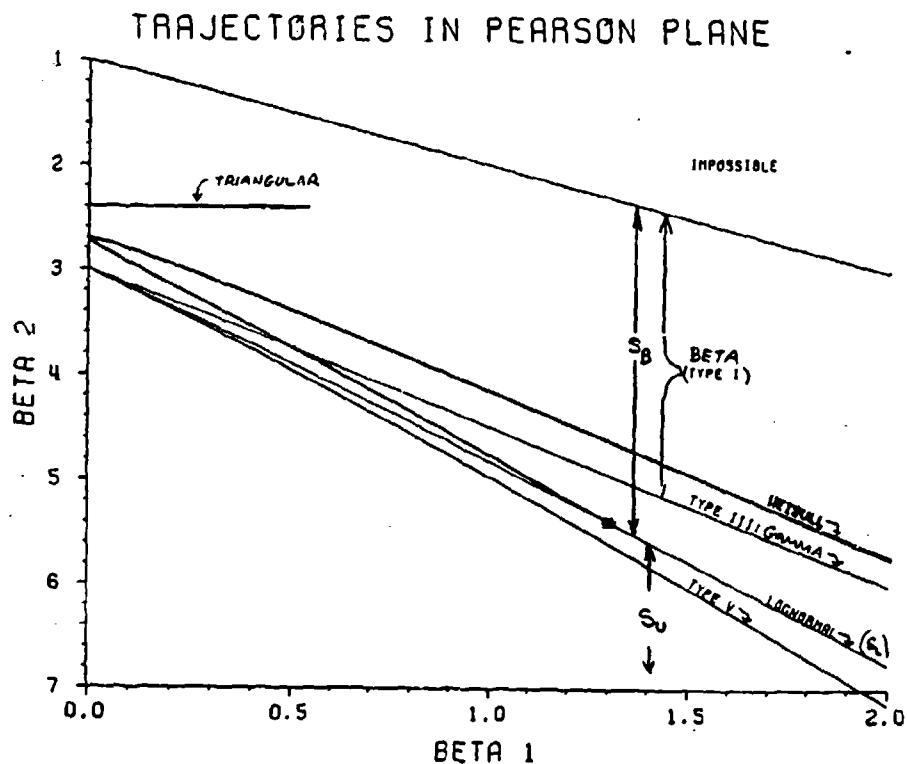
FIGURE 2-6



Percentile information is valuable for checking consistency. If figure 2-1 accurately captures the analyst's uncertainty about larger costs, the odds against the cost exceeding 27.378 are 3-1. If the analyst likes the left part of the picture but feels that 3-1 odds are too generous, he may use option 4 to lengthen the right tail. If he is uncomfortable with giving or taking even odds that the cost will not exceed 22.192, then some sort of shift may be indicated. Including percentile information would enhance the graphical selection scheme proposed in [9].

There is no scientific reason for using beta distributions (or any other, for that matter) in quantifying cost uncertainty. Arguments supporting other choices (e.g. triangular [14], Weibull [8]) are usually based on ease of computation. In this day of the inexpensive and powerful microcomputer, computational ease at the expense of flexibility of application seems like a poor tradeoff. The main criterion for judging the suitability of a family of densities should be its richness of possible shapes. A particularly nice shape-rich family which seems to have been overlooked was proposed by N. L. Johnson [3] and bears his name. A clear account is given in the text [2]. The family is much richer in shape possibilities than the beta family. Its skewness (beta 1) and kurtosis (beta 2) cover the entire Pearson plane shown in figure 3 with subfamily Sb covering the region above the lognormal line and subfamily Su covering the region below. The lognormal family belongs to the Johnson family and is denoted S1. By contrast, Weibull and triangular shape parameters are restricted to the curves shown. A particularly nice feature of the Johnson family is that its percentiles can be calculated using a table of standard normal probabilities. This facilitates Monte Carlo simulation work. Research on fitting these distributions to subjective assessments is underway [12] and preliminary results are encouraging.

FIGURE 3: Skewness and Kurtosis of Several Families of Densities



2. CER UNCERTAINTY

In Section 1, three sources of uncertainty which contribute to overall CER uncertainty were identified. The task of quantifying opinion about model parameters such as a and b is necessary in order to generate an uncertainty distribution over the subsystem cost. Direct assessment of a joint distribution is particularly difficult. In a recent paper [5], a technique has been developed to indirectly assess this distribution. Opinions expressed by the analyst are treated as data to estimate the model parameters. By asking questions about the predictive distribution at various values of the independent variables, it

can be inferred what the prior distribution has to be in order to be consistent with these assessments. The elicited assessments are in the form of quantile of the predictive distribution. This approach was taken in the belief that quantiles are "less subject to erratic behavior, easier to think about, and consequently easier to elicit." Consistency checks are made in the process so that answers that appear to be out of line with other answers can be called to the attention of the user. The techniques have been implemented with an interactive program on the TROLL experimental system operated by MIT. With proper training, this new approach may become a useful tool in cost uncertainty analysis.

BIBLIOGRAPHY

- [1] Fisher, G. H. (1971) Cost Considerations in Systems Analysis, American Elsevier Publishing Company, New York
- [2] Hahn, G. J. and Shapiro, S. S. (1967) Statistical Models in Engineering John Wiley and Sons, Inc, New York
- [3] Johnson, N. L. (1949) Systems of frequency curves generated by methods of translation. *Biometrika*, 36, 249
- [4] Jordan, H. R. and Klein, M. R. (1975) An application of subjective probabilities to the problem of uncertainty in cost analysis, Resource Analysis Group, CNO, U.S. Navy
- [5] Kadane, J. B., Dickey, J. M., Winkler, R. L., Smith, W. S., and Peters, S. C. (1980) Interactive elicitation of opinion for a normal linear model, *Journal of the American Statistical Association*, 75, pp. 845-855
- [6] Kahneman, D., Slovic, P. and Tversky, A. (1982) Judgement Under Uncertainty: Heuristics and Biases, Cambridge University Press, New York

- [7] McNichols, G. R. (1976) On the treatment of uncertainty in parametric costing, Report T-330, Inst. of Management Science and Engineering, The George Washington Univ., Washington, D. C.
- [8] McNichols, G. R. (1979) Treatment of uncertainty in life cycle costing, Proceedings of IEEE Symposium on Reliability and Maintainability
- [9] McNichols, G. R. (1984) The state-of-the-art of cost uncertainty analysis, The Journal of Cost Analysis, 1, pp. 149-174
- [10] Wallenius, K. T. (1983) Some observation on a subjective probability estimation technique (SPET) TR 437, Dept. of Mathematical Sciences, Clemson University.
- [11] Wallenius, K. T. (1985) Cost uncertainty analysis methodology: A critical overview, Proceedings, 19th Annual DOD Cost Analysis Symposium
- [12] Wallenius, K. T. (1985) The Johnson distributions as uncertainty quantifiers, TR 506, Dept. of Mathematical Sciences, Clemson University.
- [13] Wilder, J. J. and Black R .(1981) Using moments in cost risk analysis, Proceedings of Management of Risk and Uncertainty Symposium, U. S. Air Force Academy, Colorado Springs
- [14] Wilder, J. J. (1973) A simplified approach to Monte Carlo cost risk evaluation, Grumman Aerospace Corp. PDR-OP-T73-27

END

FILMED

1-86

DTIC